

OIPPE

RAW SEQUENCE LISTING

DATE: 08/15/2001

PATENT APPLICATION: US/09/530,233

TIME: 08:21:43

Input Set : A:\Mcgill.app

Output Set: N:\CRF3\08132001\I530233.raw

3 <110> APPLICANT: Seguela, Philippe
 4 Babinski, Kazimierz
 5 McGill University
 7 <120> TITLE OF INVENTION: DNA ENCODING A HUMAN PROTON-GATED ION CHANNEL AND USES
 8 THEREOF
 10 <130> FILE REFERENCE: 641050.90021
 12 <140> CURRENT APPLICATION NUMBER: US/09/530,233
 13 <141> CURRENT FILING DATE: 1997-10-29
 15 <150> PRIOR APPLICATION NUMBER: US09/530,233
 16 <151> PRIOR FILING DATE: 1997-10-29
 18 <160> NUMBER OF SEQ ID NOS: 5
 20 <170> SOFTWARE: PatentIn Ver. 2.0
 22 <210> SEQ ID NO: 1
 23 <211> LENGTH: 1732
 24 <212> TYPE: DNA
 25 <213> ORGANISM: Homo sapiens
 27 <220> FEATURE:
 28 <221> NAME/KEY: CDS
 29 <222> LOCATION: (22)..(1614)
 31 <400> SEQUENCE: 1

| | | | | | | | | | | | | | | |
|----|------------|-------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 32 | tcgcacgacg | cgggttctggc | c | atg | aag | ccc | acc | tca | ggc | cca | gag | gag | gcc | 51 |
| 33 | | | | Met | Lys | Pro | Thr | Ser | Gly | Pro | Glu | Glu | Ala | |
| 34 | | | | 1 | | | | 5 | | | | | 10 | |
| 36 | cgg | cgg | cag | ccc | tcg | gac | atc | cgc | gtg | ttc | gcc | agc | aac | 99 |
| 37 | Arg | Arg | Gln | Pro | Ser | Asp | Ile | Arg | Val | Phe | Ala | Ser | Asn | |
| 38 | | | | 15 | | | | 20 | | | | | 25 | |
| 40 | cac | ggg | ctg | ggc | cac | gtc | ttc | ggg | cca | ggc | agc | ctg | agc | 147 |
| 41 | His | Gly | Leu | Gly | His | Val | Phe | Gly | Pro | Gly | Ser | Leu | Ser | |
| 42 | | | | 30 | | | | 35 | | | | | 40 | |
| 44 | ggg | atg | tgg | gca | gcg | gcc | gtg | gtc | ctg | tca | gtg | gcc | acc | 195 |
| 45 | Gly | Met | Trp | Ala | Ala | Ala | Val | Val | Leu | Ser | Val | Ala | Thr | |
| 46 | | | | 45 | | | | 50 | | | | | 55 | |
| 48 | cag | gtg | gct | gag | agg | gtg | cgc | tac | tac | agg | gag | ttc | cac | 243 |
| 49 | Gln | Val | Ala | Glu | Arg | Val | Arg | Tyr | Tyr | Arg | Glu | Phe | His | |
| 50 | | | | 60 | | | | 65 | | | | | 70 | |
| 52 | gcc | ctg | gat | gag | cga | gaa | agc | cac | cgg | ctc | gtc | ttc | ccg | 291 |
| 53 | Ala | Leu | Asp | Glu | Arg | Glu | Ser | His | Arg | Leu | Val | Phe | Pro | |
| 54 | | | | 75 | | | | 80 | | | | | 85 | |
| 56 | ctg | tgc | aac | atc | aac | cca | ctg | cgc | cgc | tcg | cgc | cta | acg | 339 |
| 57 | Leu | Cys | Asn | Ile | Asn | Pro | Leu | Arg | Arg | Ser | Arg | Leu | Thr | |
| 58 | | | | 95 | | | | 100 | | | | | 105 | |
| 60 | ctg | cac | tgg | gct | ggg | tct | gcg | ctg | ctg | ggc | ctg | gat | ccc | 387 |
| 61 | Leu | His | Trp | Ala | Gly | Ser | Ala | Leu | Leu | Gly | Leu | Asp | Pro | |
| 62 | | | | 110 | | | | 115 | | | | | 120 | |
| 64 | gcc | gcc | ttc | ctg | cgc | gcc | ctg | ggc | cgg | ccc | cct | gca | ccg | 435 |
| 65 | Ala | Ala | Phe | Leu | Arg | Ala | Leu | Gly | Arg | Pro | Pro | Ala | Pro | |
| 66 | | | | 125 | | | | 130 | | | | | 135 | |

ENTERED

RAW SEQUENCE LISTING

DATE: 08/15/2001

PATENT APPLICATION: US/09/530,233

TIME: 08:21:43

Input Set : A:\Mcgill.app

Output Set: N:\CRF3\08132001\I530233.raw

```

68 atg ccc agt ccc acc ttt gac atg gcg caa ctc tat gcc cgt gct ggg 483
69 Met Pro Ser Pro Thr Phe Asp Met Ala Gln Leu Tyr Ala Arg Ala Gly
70 140 145 150
72 cac tcc ctg gat gac atg ctg ctg gac tgt cgc ttc cgt ggc caa cct 531
73 His Ser Leu Asp Asp Met Leu Leu Asp Cys Arg Phe Arg Gly Gln Pro
74 155 160 165 170
76 tgt ggg cct gag aac ttc acc acg atc ttc acc cgg atg gga aag tgc 579
77 Cys Gly Pro Glu Asn Phe Thr Thr Ile Phe Thr Arg Met Gly Lys Cys
78 175 180 185
80 tac aca ttt aac tct ggc gct gat ggg gca gag ctg ctc acc act act 627
81 Tyr Thr Phe Asn Ser Gly Ala Asp Gly Ala Glu Leu Leu Thr Thr Thr
82 190 195 200
84 agg ggt ggc atg ggc aat ggg ctg gac atc atg ctg gac gtg cag cag 675
85 Arg Gly Gly Met Gly Asn Gly Leu Asp Ile Met Leu Asp Val Gln Gln
86 205 210 215
88 gag gaa tat cta cct gtg tgg agg gac aat gag gag acc ccg ttt gag 723
89 Glu Glu Tyr Leu Pro Val Trp Arg Asp Asn Glu Glu Thr Pro Phe Glu
90 220 225 230
92 gtg ggg atc cga gtg cag atc cac agc cag gag gag ccg ccc atc atc 771
93 Val Gly Ile Arg Val Gln Ile His Ser Gln Glu Glu Pro Pro Ile Ile
94 235 240 245 250
96 gat cag ctg ggc ttg ggg gtg tcc ccg ggc tac cag acc ttt gtt tct 819
97 Asp Gln Leu Gly Leu Gly Val Ser Pro Gly Tyr Gln Thr Phe Val Ser
98 255 260 265
100 tgc cag cag cag cag ctg agc ttc ctg cca ccg ccc tgg ggc gat tgc 867
101 Cys Gln Gln Gln Gln Leu Ser Phe Leu Pro Pro Pro Trp Gly Asp Cys
102 270 275 280
104 agt tca gca tct ctg aac ccc aac tat gag cca gag ccc tct gat ccc 915
105 Ser Ser Ala Ser Leu Asn Pro Asn Tyr Glu Pro Glu Pro Ser Asp Pro
106 285 290 295
108 cta ggc tcc ccc agc ccc agc ccc agc cct ccc tat acc ctt atg ggg 963
109 Leu Gly Ser Pro Ser Pro Ser Pro Ser Pro Pro Tyr Thr Leu Met Gly
110 300 305 310
112 tgt cgc ctg gcc tgc gaa acc cgc tac gtg gct cgg aag tgc ggc tgc 1011
113 Cys Arg Leu Ala Cys Glu Thr Arg Tyr Val Ala Arg Lys Cys Gly Cys
114 315 320 325 330
116 cga atg gtg tac atg cca ggc gac gtg cca gtg tgc agc ccc cag cag 1059
117 Arg Met Val Tyr Met Pro Gly Asp Val Pro Val Cys Ser Pro Gln Gln
118 335 340 345
120 tac aag aac tgt gcc cac ccg gcc ata gat gcc atc ctt cgc aag gac 1107
121 Tyr Lys Asn Cys Ala His Pro Ala Ile Asp Ala Ile Leu Arg Lys Asp
122 350 355 360
124 tcg tgc gcc tgc ccc aac ccg tgc gcc agc acg cgc tac gcc aag gag 1155
125 Ser Cys Ala Cys Pro Asn Pro Cys Ala Ser Thr Arg Tyr Ala Lys Glu
126 365 370 375
128 ctc tcc atg gtg cgg atc ccg agc cgc gcc gcc gcg cgc ttc ctg gcc 1203
129 Leu Ser Met Val Arg Ile Pro Ser Arg Ala Ala Ala Arg Phe Leu Ala
130 380 385 390
132 cgg aag ctc aac cgc agc gag gcc tac atc gcg gag aac gtg ctg gcc 1251

```

RAW SEQUENCE LISTING

DATE: 08/15/2001

PATENT APPLICATION: US/09/530,233

TIME: 08:21:43

Input Set : A:\Mcgill.app

Output Set: N:\CRF3\08132001\I530233.raw

```

133 Arg Lys Leu Asn Arg Ser Glu Ala Tyr Ile Ala Glu Asn Val Leu Ala
134 395 400 405 410
136 ctg gac atc ttc ttt gag gcc ctc aac tat gag acc gtg gag cag aag 1299
137 Leu Asp Ile Phe Phe Glu Ala Leu Asn Tyr Glu Thr Val Glu Gln Lys
138 415 420 425
140 aag gcc tat gag atg tca gag ctg ctt ggt gac att ggg ggc cag atg 1347
141 Lys Ala Tyr Glu Met Ser Glu Leu Leu Gly Asp Ile Gly Gly Gln Met
142 430 435 440
144 ggc ctt ttc atc ggg gcc agc ctg ctc acc atc ctc gag atc cta gac 1395
145 Gly Leu Phe Ile Gly Ala Ser Leu Leu Thr Ile Leu Glu Ile Leu Asp
146 445 450 455
148 tac ctc tgt gag gtg ttc cga gac aag gtc ctg gga tat ttc tgg aac 1443
149 Tyr Leu Cys Glu Val Phe Arg Asp Lys Val Leu Gly Tyr Phe Trp Asn
150 460 465 470
152 cga cag cac tcc caa agg cac tcc agc acc aat ctg ctt cag gaa ggg 1491
153 Arg Gln His Ser Gln Arg His Ser Ser Thr Asn Leu Leu Gln Glu Gly
154 475 480 485 490
156 ctg ggc agc cat cga acc caa gtt ccc cac ctc agc ctg ggc ccc aga 1539
157 Leu Gly Ser His Arg Thr Gln Val Pro His Leu Ser Leu Gly Pro Arg
158 495 500 505
160 cct ccc acc cct ccc tgt gcc gtc acc aag act ctc tcc gcc tcc cac 1587
161 Pro Pro Thr Pro Pro Cys Ala Val Thr Lys Thr Leu Ser Ala Ser His
162 510 515 520
164 cgc acc tgc tac ctt gtc aca cag ctc tagacctgct gtctgtgtcc 1634
165 Arg Thr Cys Tyr Leu Val Thr Gln Leu
166 525 530
168 tcggagcccc gccctgacat cctggacatg cctagcctgc acgtagcttt tccgtcttca 1694
170 ccccaaataa agtcctaata catcaaaaaa aaaaaaaa 1732
173 <210> SEQ ID NO: 2
174 <211> LENGTH: 531
175 <212> TYPE: PRT
176 <213> ORGANISM: Homo sapiens
178 <400> SEQUENCE: 2
179 Met Lys Pro Thr Ser Gly Pro Glu Glu Ala Arg Arg Gln Pro Ser Asp
180 1 5 10 15
182 Ile Arg Val Phe Ala Ser Asn Cys Ser Met His Gly Leu Gly His Val
183 20 25 30
185 Phe Gly Pro Gly Ser Leu Ser Leu Arg Arg Gly Met Trp Ala Ala Ala
186 35 40 45
188 Val Val Leu Ser Val Ala Thr Phe Leu Tyr Gln Val Ala Glu Arg Val
189 50 55 60
191 Arg Tyr Tyr Arg Glu Phe His His Gln Thr Ala Leu Asp Glu Arg Glu
192 65 70 75 80
194 Ser His Arg Leu Val Phe Pro Ala Val Thr Leu Cys Asn Ile Asn Pro
195 85 90 95
197 Leu Arg Arg Ser Arg Leu Thr Pro Asn Asp Leu His Trp Ala Gly Ser
198 100 105 110
200 Ala Leu Leu Gly Leu Asp Pro Ala Glu His Ala Ala Phe Leu Arg Ala
201 115 120 125

```

RAW SEQUENCE LISTING

DATE: 08/15/2001

PATENT APPLICATION: US/09/530,233

TIME: 08:21:43

Input Set : A:\Mcgill.app

Output Set: N:\CRF3\08132001\I530233.raw

```

203 Leu Gly Arg Pro Pro Ala Pro Pro Gly Phe Met Pro Ser Pro Thr Phe
204      130      135      140
206 Asp Met Ala Gln Leu Tyr Ala Arg Ala Gly His Ser Leu Asp Asp Met
207 145      150      155      160
209 Leu Leu Asp Cys Arg Phe Arg Gly Gln Pro Cys Gly Pro Glu Asn Phe
210      165      170      175
212 Thr Thr Ile Phe Thr Arg Met Gly Lys Cys Tyr Thr Phe Asn Ser Gly
213      180      185      190
215 Ala Asp Gly Ala Glu Leu Leu Thr Thr Thr Arg Gly Gly Met Gly Asn
216      195      200      205
218 Gly Leu Asp Ile Met Leu Asp Val Gln Gln Glu Glu Tyr Leu Pro Val
219      210      215      220
221 Trp Arg Asp Asn Glu Glu Thr Pro Phe Glu Val Gly Ile Arg Val Gln
222 225      230      235      240
224 Ile His Ser Gln Glu Glu Pro Pro Ile Ile Asp Gln Leu Gly Leu Gly
225      245      250      255
227 Val Ser Pro Gly Tyr Gln Thr Phe Val Ser Cys Gln Gln Gln Gln Leu
228      260      265      270
230 Ser Phe Leu Pro Pro Pro Trp Gly Asp Cys Ser Ser Ala Ser Leu Asn
231      275      280      285
233 Pro Asn Tyr Glu Pro Glu Pro Ser Asp Pro Leu Gly Ser Pro Ser Pro
234      290      295      300
236 Ser Pro Ser Pro Pro Tyr Thr Leu Met Gly Cys Arg Leu Ala Cys Glu
237 305      310      315      320
239 Thr Arg Tyr Val Ala Arg Lys Cys Gly Cys Arg Met Val Tyr Met Pro
240      325      330      335
242 Gly Asp Val Pro Val Cys Ser Pro Gln Gln Tyr Lys Asn Cys Ala His
243      340      345      350
245 Pro Ala Ile Asp Ala Ile Leu Arg Lys Asp Ser Cys Ala Cys Pro Asn
246      355      360      365
248 Pro Cys Ala Ser Thr Arg Tyr Ala Lys Glu Leu Ser Met Val Arg Ile
249      370      375      380
251 Pro Ser Arg Ala Ala Ala Arg Phe Leu Ala Arg Lys Leu Asn Arg Ser
252 385      390      395      400
254 Glu Ala Tyr Ile Ala Glu Asn Val Leu Ala Leu Asp Ile Phe Phe Glu
255      405      410      415
257 Ala Leu Asn Tyr Glu Thr Val Glu Gln Lys Lys Ala Tyr Glu Met Ser
258      420      425      430
260 Glu Leu Leu Gly Asp Ile Gly Gly Gln Met Gly Leu Phe Ile Gly Ala
261      435      440      445
263 Ser Leu Leu Thr Ile Leu Glu Ile Leu Asp Tyr Leu Cys Glu Val Phe
264      450      455      460
266 Arg Asp Lys Val Leu Gly Tyr Phe Trp Asn Arg Gln His Ser Gln Arg
267 465      470      475      480
269 His Ser Ser Thr Asn Leu Leu Gln Glu Gly Leu Gly Ser His Arg Thr
270      485      490      495
272 Gln Val Pro His Leu Ser Leu Gly Pro Arg Pro Pro Thr Pro Pro Cys
273      500      505      510
275 Ala Val Thr Lys Thr Leu Ser Ala Ser His Arg Thr Cys Tyr Leu Val

```

RAW SEQUENCE LISTING

DATE: 08/15/2001

PATENT APPLICATION: US/09/530,233

TIME: 08:21:43

Input Set : A:\Mcgill.app

Output Set: N:\CRF3\08132001\I530233.raw

```

276          515          520          525
278 Thr Gln Leu
279          530
282 <210> SEQ ID NO: 3
283 <211> LENGTH: 17
284 <212> TYPE: PRT
285 <213> ORGANISM: Artificial Sequence ✓
287 <220> FEATURE:
288 <221> NAME/KEY: PEPTIDE
289 <222> LOCATION: (2)
290 <223> OTHER INFORMATION: "Xaa = unidentified amino acid."
292 <220> FEATURE:
293 <221> NAME/KEY: PEPTIDE
294 <222> LOCATION: (11)
295 <223> OTHER INFORMATION: "Xaa = unidentified amino acid."
297 <220> FEATURE:
298 <221> NAME/KEY: PEPTIDE
299 <222> LOCATION: (13)..(14)
300 <223> OTHER INFORMATION: "Xaa = unidentified amino acid."
302 <220> FEATURE:
303 <221> NAME/KEY: PEPTIDE
304 <222> LOCATION: (16)
305 <223> OTHER INFORMATION: Xaa = unidentified amino acid."
307 <220> FEATURE:
308 <223> OTHER INFORMATION: Description of Artificial Sequence: Protein Motif ✓
310 <400> SEQUENCE: 3
W--> 311 Leu Xaa Phe Pro Ala Val Thr Leu Cys Asn Xaa Asn Xaa Xaa Arg Xaa
312      1              5              10              15
314 Ser
318 <210> SEQ ID NO: 4
319 <211> LENGTH: 20
320 <212> TYPE: DNA
321 <213> ORGANISM: Artificial Sequence ✓
323 <220> FEATURE:
324 <223> OTHER INFORMATION: Description of Artificial Sequence: ✓
325      Oligonucleotide Primer
327 <400> SEQUENCE: 4
328 tcagtggcca ccttcctcta
331 <210> SEQ ID NO: 5
332 <211> LENGTH: 21
333 <212> TYPE: DNA
334 <213> ORGANISM: Artificial Sequence ✓
336 <220> FEATURE:
337 <223> OTHER INFORMATION: Description of Artificial Sequence: ✓
338      Oligonucleotide Primer
340 <400> SEQUENCE: 5
341 acagtccagc agcatgtcat c

```

20

21

VERIFICATION SUMMARY

DATE: 08/15/2001

PATENT APPLICATION: US/09/530,233

TIME: 08:21:44

Input Set : A:\Mcgill.app

Output Set: N:\CRF3\08132001\I530233.raw

L:12 M:270 C: Current Application Number differs, Replaced Current Application Number

L:13 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:311 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3